

ROBUST DOCUMENT BOUNDARY DETERMINATION

BACKGROUND OF THE INVENTION

The present application relates to document boundary determination.

5 Optical scanners operate by imaging an object, typically in the form of a sheet of paper, document, or other form of medium with a light source. The optical scanner senses a resultant light signal from the medium with an optical sensor array that includes pixel elements generating a data signal representative of the intensity of light impinging thereon for a corresponding portion of the imaged object. The data signals
10 from the array are then processed (typically digitized) and utilized by a utilization apparatus or stored on a suitable medium such as a hard drive of a computer system for subsequent display and/or manipulation.

 Various types of photo sensor devices may be used in optical scanners. For example, a commonly used photo sensor device is the charge coupled device (CCD),
15 which builds up an electrical charge in response to exposure to light. The size of the electrical charge built up is dependent on the intensity and duration of the light exposure. In optical scanners, CCD cells are typically arranged in linear arrays. Each cell or "pixel" has a portion of a scan line image impinged thereon as the scan line sweeps across the scanned object. The charge build up in each of the pixels is measured and discharged at
20 regular "sampling" intervals.

 The image of a scan line portion of a document is projected onto the scanner's linear sensor array by scanner optics. In CCD scanners, the scanner optics typically comprise an imaging lens which typically reduces the size of the projected

5 image from the original size of the document. Pixels in a scanner linear photo sensor array are aligned in a direction perpendicular to the "scan" direction, i.e. the paper or scanner movement direction for scanning of the image.

At any instant when an object is being scanned, each pixel in the sensor array has a corresponding area on the object which is being imaged thereon. This
10 corresponding area on the scanned object is referred to as an "object pixel." An area on a scanned object corresponding in area to the entire area of the linear sensor array is referred to as an "object scan line" or "scan line." For descriptive purposes, a scanned object is considered to have a series of fixed adjacently positioned scan lines. Scanners are typically operated at a scan line sweep rate such that one scan line width is traversed
15 during each sampling interval.

Some optical scanning machines include an automatic document feeder for feeding a document past the optical array. Other optical scanners machines are known as "flat-bed" scanners, wherein a document is placed on a fixed platen for scanning, which occurs by moving the sensor array relative to the fixed document.

20 It is advantageous in various applications to sense the location of a document edge. In a printer, for example, the print area differs depending on whether the printing is on envelopes, name card paper, letter sized paper, and so on. The prediction of the print area assists in driving the print head. The print area can be identified by sensing the media edges. By identifying the document area, proper clipping can be made on both
25 sides when printing. In a scanner, detection of the document edges can assist by placing the image area properly on the page, and by reducing the scan memory size by clipping the empty regions. Also, by detecting the edge position in the direction of document

5 movement, the document skew can be estimated and used to redirect the scanned image
in print. This will produce a more pleasing output from the scanning process. In a
copier, sensing the size of a document permits scaling of the input document to the
maximum size that will fit on an output page. In addition, multi-function machines
combine in a single machine the functions of printing and optical scanning with
10 automatic document/sheet feeders.

If a document is misaligned with respect to the optical sensor, the resultant
image is similarly skewed. Because the contents of a document page are usually aligned
with the page itself, a skewed page usually results in a misalignment with the optical
sensor.

15 Pasco et al., U.S. Patent No. 5,818,976, disclose a system for skew and
size/shape detection of a document. The system performs the following basic steps, (1)
detects points near the edge of the page image, (2) fits lines to establish a closed contour,
and (3) defines a polygon with sides coincident with the lines of the closed contour. The
polygon defines the size and shape of the page image. With respect to detecting the
20 edges of the page the system uses a background (platen backing cover) that contrasts well
with the page, e.g., a black (or gray or patterned) background and white documents. Then
the system analyzes the image to determine the edges of the image. Unfortunately, this
requires specialized hardware to determine the edge of the image and thus is unsuitable
for general purpose scanning devices. If a contrasting background is not used, Pasco et
25 al. suggest the use of electro-mechanical switches or optical switches arranged to sense
the location of edges of each page in conjunction with scanning. Likewise, this requires

5 specialized hardware to determine the edge of the image and thus is unsuitable for general purpose scanning devices.

What is desired, therefore, is a system that can determine the general bounding region of a document without additional specialized hardware.

10 DETAILED DESCRIPTION OF THE DRAWINGS

FIG. 1 is an exemplary illustration of a scanner, document and cover.

FIGS. 2A, 2B AND 2C illustrate sample images.

FIG. 3 is an exemplary process flow chart.

FIGS. 4A and 4B illustrate stat buffers.

15 FIGS. 5A and 5B illustrate accumulation buffers.

FIGS. 6A and 6B illustrate smoothed buffers.

FIG. 7 is an exemplary example of document boundary detection of a small document.

FIG. 8 is an exemplary example of document boundary detection of FIG. 8.

20 FIG. 9 is an exemplary example of document boundary detection of a large multi-opaqueness document.

FIG. 10 is an exemplary example of document boundary detection of FIG. 9.

FIG. 11 is an exemplary neighborhood of a document boundary edge shown in FIG. 7.

25 FIG. 12 is an exemplary document boundary edge of a cluster of noise points shown in FIG. 7.

FIG. 13 is an exemplary edge determination for row stat buffers.

5 FIG. 14 is an exemplary edge determination for column stat buffers.

FIG. 15 is an exemplary example of proper document boundary detection for the small document down in FIG. 7.

FIG. 16 is an exemplary example of proper document boundary detection for the large document shown in FIG. 9.

10

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENT

15 The present inventors considered the existing prior art systems, which generally use specialized physical devices such as sheet feeders, document delivery systems, specially designed platen covers, multiple light sources, etc. Each of the systems is unsuitable for general purpose document edge detection because it requires modification or otherwise specialized design of the hardware for the system. The present inventors then further considered typical existing scanning devices and came to the realization that many include a cover thereon under which the original document is positioned. Traditional wisdom suggests that a cover having substantially the same color as the background of the document contained thereunder, such as a white colored cover and a white document, is unsatisfactory for determining the edge of the document. For example, Pasco. et al., U.S. Patent No. 5,818,976, in fact teach the use of a cover that contrasts well with the page. In direct contrast to this traditional wisdom, the present inventors realized that a cover with the same general color as the document itself may be used in determining the boundary of a document. In actual systems, the document does not tend to lay perfectly flush against the cover and accordingly, when the document is

20

25

5 illuminated, a slight shadow is cast by the document onto the cover along a sufficient portion of the edge of the document.

Referring to FIG. 1, a document 10 is supported by a scanning device 12 with a cover 14 covering the document 10, all of which may be flat. Preferably, the cover is substantially flat. Preferably, a major portion of or all of, the cover has substantially
10 the same color as the background color of the document to be scanned. More specifically, preferably a major portion of the portion of the cover proximate the edge of the document preferably has substantially the same color as the background color of the document to be scanned. The scanning device may be any type of device capable of obtaining or otherwise sensing an image of the document 10. The document 10 may be
15 any type of document or otherwise an object that is sensed by the scanning device 12. In addition, the scanner may use a roller or other backing arranged in a manner opposing the imaging system with respect to the document.

Referring also to FIGS. 2A-2C, the document 10 including a portion of the cover 14 extending beyond the periphery of the document 10 is imaged or otherwise
20 sensed by the system. FIG. 2A illustrates a scanned document with a skew, together with horizontal and vertical boundary lines. FIG. 2B illustrates a scanned document with wrinkles, together with horizontal and vertical boundary lines. FIG. 2C illustrates a small document, together with horizontal and vertical boundary lines. The image acquisition may be a normal scan, a preview scan at a lower resolution than the normal scan, a
25 preview scan at a higher resolution than the normal scan or any other type of image acquisition. The image is normally acquired in a color space that includes red, green, and blue. Alternatively, any set of one or more colors may be used, black and white, or any

5 other image description scheme. The edges of the document 10 cast a slight shadow onto the cover, at least a portion of which is likewise sensed. The resulting image together with the shadow may be processed in any suitable manner to determine the size or boundaries of the document.

The process described below is suitable for processing image documents
10 in general and particularly suitable for processing the imaged document shown in FIGS. 2A-2C. The process is defined in terms of the document being properly registered with respect to the top-left corner of the platen. It is to be understood that the process may be readily extended to the document being located at any position, including a random position and random orientation. In addition, the number (e.g., one, two, three, four, five,
15 etc) of vertical and horizontal boundary lines (including other orientations of the boundary lines, such as inclined) may be extended depending on the location and shape of the document. Referring to FIG. 3, after acquisition of the image 20 in a red, blue, green color space the image 20 is preferably converted from the red, green, blue color space to a color space that enhances the luminance of the image at block 40. With a
20 document that is sensed under relatively uniform illumination, especially when narrow shadows are to be sensed, it is preferable to process the image therefrom in terms of enhanced luminance. It is to be understood that the image may be processed in any color space, as desired. The conversion from the red, green, blue color space to luminance Y may be computed as: $Y=(0.3R +0.59G + 0.11B)$. Preferably the acquired image 20 is
25 obtained at a lower resolution than the normal resolution used by the system for creating a copy in order to reduce the memory requirements. The conversion from a triplet color space (e.g., red, green, blue) to a luminance results in a reduction of the data by

5 approximately a third, which reduces the memory requirements of the system and the computational complexity.

Preferably after converting the image to a color space that enhances luminance, the predicted range of values representative of anticipated document boundary edges may be stretched or otherwise enhanced to provide a greater weight, sensitivity, or
10 otherwise, at block 50. Stretching increases the robustness of the edge detection process and enhances shadow edges by increasing the differences of pixel values in the range of likely document edge values and by attenuating edge magnitudes in the color range of the scanner cover and other data such as text. For example, pixels having a luminance in the range of 190-220 (potential values from 0-256) may be stretched to the range of 170-240
15 by applying an S-curve. It is to be understood that any modification of the image to enhance image characteristics likely to be characteristic of the edge of a document may be used. In addition, the image modification by conversion to luminance enhancement, stretching, if performed at all, may be performed at any time during processing.

Preferably after converting and stretching the image, the image is down
20 sampled to a lower resolution, such as 75 x 75 dpi, at block 60. Down sampling the image from a 300x150 preview scan resolution to a 75 x 75 resolution results in approximately an 8 times reduction in the data. This likewise results in a consistent sampling density for further processing, which provides greater consistency for image processing and more flexibility in implementing the system on different platforms. A 75
25 x 75 resolution generally results in no more than 640 x 896 pixels (an A4 U.S. letter-sized scanner platen is assumed without loss of generality). In addition, down sampling the luminance enhanced data is less computationally intensive than down sampling the

5 original image data. For example, a 1 x 4 box filter average in the horizontal direction and a two tap IIR filter in the vertical direction may be used.

The image resulting from the down sampling 60 may subsequently be divided into row strips (e.g., 32-pixels high) and column strips (e.g., 32-pixels wide). For each row strip, a set of contiguous sub-rows may be selected, such as 8, 16, or 32 rows. 10 For each column strip, a set of contiguous sub-columns may be selected, such as 8, 16, or 32 columns. In essence, the down sampled image 60 is partitioned into a set horizontal strips consisting of a group of rows, and into a set of vertical strips consisting of a group of columns. It is to be understood that any number of pixel strips, any number of sub-strips, contiguous or non-contiguous, may be used. For the illustrated example, 8 element 15 sub-strips are used, simply for ease of illustration.

The transverse average for each horizontal sub-strip is computed and stored in a horizontal stat buffer, at block 70. It is to be understood that any other statistical measure for each horizontal sub-strip may likewise be used, as desired. In the particular example illustrated, each 8-row horizontal sub-row is 640 columns wide. In 20 the illustrated example there are 28 such 640 element row-stat buffers for the image. An exemplary row stat buffer is shown in FIG. 4A for the image shown in FIG. 2A.

Similarly, the transverse average for each vertical sub-strip is computed and stored in a vertical stat buffer, at block 80. It is to be understood that any other statistical measure for each vertical sub-strip may likewise be used, as desired. In the 25 particular example illustrated each 8-column sub-strip is 896 rows high. In this example there are 20 such 896 element column stat buffers. The number and length of the row and

5 column buffers may be selected, as desired. An exemplary column stat buffer is shown in FIG. 4B for the image shown in FIG. 2A.

The use of column and row statistical buffers permits the simulation of a larger convolution kernel which results in more robust processing and likewise reduces the amount of data. Further, the transverse processing reinforces the image detail in a transverse direction which emphasizes the shadow on the edges of a document. In addition, the relatively tall filter relative to the typical height of the text tends to attenuate the text.

At block 90 a localized 1-dimensional difference operator identifies the edges of the image whose magnitude difference is above a selected threshold. For example, points whose measured local difference along the row (or column) stat buffer is greater than 10 may be considered edges. It is to be understood that any one or multi-dimensional operator which identifies edges in an image may likewise be used. The use of an appropriate operator tends to identify those regions of the image that are candidate regions of the shadow cast by the document. Conceptually, this results in another row stat data structure and another column data structure where edges are identified.

With the potential edge regions of the row stat buffer identified, or otherwise potential edge regions of the image, the total number of potential identified edge features for each transverse column are summed together. The total number of edge features for each transverse column is stored in an accumulation row buffer at block 100. An exemplary accumulation row buffer is illustrated in FIG. 5A, with each vertical line representative of a region of 32 pixels. Similarly, with the potential edge regions of the column stat buffer identified, or otherwise potential edge regions of the image, the total

5 number of potential identified edge features for each transverse row are summed. The
total number of edge features for each transverse row is stored in an accumulation column
buffer at block 110. An exemplary accumulation column buffer is illustrated in FIG.
5B, with each horizontal line representative of a region of 32 pixels. This results in an
increased likelihood of accurately determining the appropriate horizontal and vertical
10 positioning of the edges resulting from the cast shadow. Further, another significant
reduction of the amount of data is accomplished, e.g., 28 (rows) x 640 (columns) to 1
(row) x 640 (columns).

While transverse accumulation aids in identification of those regions of
potential shadows, however, the potential skew of the document itself tends to spatially
15 spread the apparent edge. To compensate for the potentially skewed edge, the data in the
accumulator is passed through a smoothing function at block 120, such as for example, a
Gaussian filter [1, 2, 1]. In essence, each particular value is adjusted in accordance with
its neighboring values. The effect is to emphasize values in regions having significant
spatially adjacent or proximate values. An exemplary smoothed data set of the row
20 buffer is shown in FIG. 6A and an exemplary smoothed data set of the column buffer is
shown in FIG. 6B. Alternatively, emphasizing values in regions spatially adjacent or
proximate to one another may be undertaken during other processes, such as the
accumulation process.

The boundary of the image, or otherwise the document, may be
25 determined based upon the largest value in the accumulator or as a result of the
smoothing. Another technique to determine the boundary of the image is to select the
outermost value greater than a sufficient threshold at block 130. In addition, the system

5 may likewise determine the boundary region of the image, text or other content on the document. However, larger images tend to have larger smoothed accumulator values, while smaller images tend to have smaller smoothed accumulator values. This difference in the maximum values tends to make it difficult to select an appropriate threshold value. To overcome the thresholding dilemma, the present system may incorporate a threshold
10 that is expressed as a percentage (or other statistical measure) of the maximum observed row or column edge count (or other criteria). This permits a single threshold to be used for both the horizontal and vertical boundaries, even with different sensitivities in each direction. The horizontal boundary of the document may then be considered as the right-most row edge count above the row-scaled threshold. The vertical boundary threshold
15 may then be considered as the bottom-most column edge count above the column-scaled threshold. Moreover, a single threshold value is likewise generally scale and directionally invariant.

While the aforementioned system is entirely suitable for many document imaging applications, the present inventors were surprised that the system is susceptible
20 to false positives caused by particularly large or dark pieces of dust, dirt, and other noise sources. Referring to FIG. 7, an exemplary imaging application includes a postage stamp 200 in the upper-left corner and a cluster of noise points 202 offset from the postage stamp 200 (circled for clarity purposes only). The system, when attempting to determine the outer boundary of the document, may determine the noise is the outer boundary of the
25 document and identifies region 204 as the document, as shown in FIG. 8. If the sensitivity of the system is reduced to avoid the identification of region 204 as the document, then other limitations arise as illustrated by FIG. 9. Referring to FIG. 9, an

5 exemplary imaging application includes an original document 210 that is pasted onto a thin sheet of transparent backing paper 212. The system, when attempting to determine the outer boundary of the document, typically passes over the transparent backing paper 212, and improperly determines that original document 210 as the proper boundary of the document, as shown in FIG. 10, especially if the sensitivity is reduced.

10 The principal source of the false boundary detection is the “noise data” that appears in the stat buffer. The additional unwanted data in the stat buffer becomes significant, especially if the additional detected noise points are primarily within the sampled portion of the rows and columns. A potential solution to the aforementioned dilemma is to select an appropriate threshold value that filters out the noise points while
15 still identifying the appropriate faint edges. However, the selection of a suitable fixed value is problematic.

After further consideration, the previously described system may include a variable threshold that is based upon a percentage of the maximum observed row or column edge count. With small documents, the threshold is small, which corresponds
20 with increasing the gain. Therefore when detecting small documents, small noise becomes significant and tends to result in false positives. In contrast with large documents, the threshold is large, which corresponds to decreasing the gain. Therefore when detecting large documents, small noise becomes insignificant and the system may miss faint edges.

25 In light of the foregoing limitations and the desire to accurately determine the edge boundaries of a document, the present inventors hypothesized that documents, even if skewed or otherwise, extend over a spatial extent so that edge data should exist in

5 a plurality of adjacent stat buffers. This is in contrast to attempting to solely refine the threshold technique to accommodate different sized documents. Therefore, an improved image boundary detection technique should include a spatial edge coherence determination for the data, such as the stat buffers. Accordingly, data that is non-spatially coherent in a particular direction will be ignored, or otherwise processed so as to preclude a false-positive. In other words, the present inventors came to the realization that document boundaries are usually weak in terms of edge magnitude but have a significant spatial extent of the weak edge magnitudes, while noise points may exhibit significant edge magnitudes but typically have limited spatial extent.

15 An exemplary illustration of the spatial extent of the data within the row buffers is shown in FIGS. 11 and 12. FIG. 11 shows the neighborhood of a portion of the document boundary edge of FIG. 7. The edge 230 has vertically adjacent neighbors 232. The vertically adjacent neighbors 232 indicate the existence of a potential boundary of a document, while the horizontal neighbors are typically ignored because the row buffers are designed to identify vertically oriented edges. In addition, data without a vertically adjacent neighbor may be reset to zero, or otherwise not used in the determination of the document boundary. FIG. 12 shows the neighborhood of a portion of the document boundary edge of FIG. 7, namely, that portion proximate the noise points 202. There are no vertically adjacent neighbors to the data points, therefore the data may be reset to zero, or otherwise not used in the determination of the document boundary. Alternatively, data without any vertically adjacent neighbors may be attenuated (or otherwise modified) to provide a lessened indication of a document boundary edge than it would have had the data not otherwise been attenuated (or otherwise modified). Moreover, the determination

5 of a vertically adjacent edge may be extended to edges other than those immediately adjacent, such as those a plurality of row buffers distant, and edges that include three or more positive edge determinations. Further, the determination of vertical neighbors may be extended to those offset from direct vertical positions. Moreover, the spatial directional determination may be performed at any step in the image boundary determination process, as desired. In addition, it is to be understood that this technique may likewise be used for the horizontal boundary determination.

Referring to FIGS. 13 and 14, one particular implementation of the present system is to perform the morphological test to identify and erase edge points within the row and column stat buffers. For example, each edge point that does not have a corresponding neighbor, such as a directly adjacent neighbor, in the predetermined directions is eliminated as a potential point. The elimination may be performed by replacing the edge identification with data indicative of a non-edge region, such as zero. Referring to FIG. 13, for example, row buffer region 240 may check row buffer regions 242, 244, 246, 248, 250, and 252 for adjacent neighbors to indicate whether a boundary region exists. Such a positive determination may be based on one or more of the row buffer regions 242, 244, 246, 248, 250, and 252 likewise being indicated as a potential boundary region. Referring to FIG. 14, similarly, a column buffer region 260 may check for column buffer regions 262, 264, 266, 268, 270, and 272, for adjacent neighbors to determine whether a region exists. Such a positive determination may be based on one or more of the column buffer regions 262, 264, 266, 268, 270, and 272 likewise being indicated as a potential region.

5 Referring to FIG. 15, the computed results of FIG. 7 show that the cluster
of noise points 202 have been ignored and that the right-most and bottom-most
boundaries of the postage stamp 200 have been accurately detected. Referring to FIG. 16,
the boundary of the thin backing sheet is correctly identified, despite the fact that its
boundary shadow is much fainter than the physical boundary of the document pasted on
10 top of it. At least in part, the increased robustness may result from the fact that noise
points are explicitly identified and removed. Therefore, the sensitivity of the boundary
detection process can be increased so as to detect faint boundaries without incurring false-
positive responses.

15 The terms and expressions which have been employed in the foregoing
specification are used therein as terms of description and not of limitation, and there is no
intention, in the use of such terms and expressions, of excluding equivalents of the
features shown and described or portions thereof, it being recognized that the scope of the
invention is defined and limited only by the claims which follow.

20